

Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1. (Currently amended) A method performed by a computer system, the method comprising:

receiving a first uniform resource locator (URL) ~~including one or more~~
~~parameters;~~

selecting one or more parameters present in the first URL;

generating a plurality of different URLs having different parameter combinations
of the one or more selected parameters;

retrieving ~~content corresponding to~~ using the first URL;

retrieving content ~~corresponding to~~ [[a]] using the plurality of different URLs
~~having different parameter combinations of the one or more parameters;~~

comparing, by a processor of the computer system, the content retrieved using the
first URL to the content retrieved using the plurality of different URLs;

identifying, based on the comparing, one of the [[a]] parameter combination
combinations, that, when present in a particular URL, results in retrieving from the
plurality of URLs that corresponds to content that is approximately the same as the
content corresponding to the first URL, the identifying being performed by the processor;
and

generating, by the processor, one or more URL rewrite rules based on the identified one of the parameter combination combinations.

2. (previously presented) The method of claim 1, where the different parameter combinations include the first URL with no parameters, the first URL with each of the one or more parameters individually, and the first URL with combinations of the one or more parameters.

3. (Currently amended) The method of claim 1, further comprising:
performing the receiving a first URL, the selecting one or more parameters present in the first URL, the generating a plurality of different URLs, the retrieving content ~~corresponding to a~~ using the first URL, the retrieving content ~~corresponding to a~~ using the plurality of URLs, the comparing the content, and the identifying [[a]] one of the parameter combination combinations, for multiple different first URLs, each first URL including the one or more parameters; and
generating the one or more URL rewrite rules for the identified one of the parameter combination combinations for each of the first URLs.

4. (Currently amended) The method of claim 3, where the rewrite rules specify that parameters that do not occur in a threshold number of the identified one of the parameter combination combinations are to be removed.

5. (previously presented) The method of claim 1, where each rewrite rule applies to a particular web site or web host.

6. (Currently amended) The method of claim 1, where the identified one of the parameter combination combinations includes a minimum number of parameters with respect other ones of the parameter combinations.

7. (Currently amended) A method, performed by a computer system, for converting a uniform resource locator (URL) into a canonical form of the URL, the method comprising:

receiving a URL that refers to content and that includes a parameter set including at least one parameter;

~~determining~~ selecting, by a processor of the computer system, a rewrite rule by receiving a plurality of URLs that include the parameter set and identifying parameters in the parameter set that do not ~~contribute to~~ result in retrieving substantially different content, when present in a URL;

applying, by the processor, the rewrite rule to the URL by removing the parameters that do not contribute to content from the URL; and

outputting the rewritten URL as the canonical form of the URL.

8. (Canceled)

9. (Currently amended) The method of claim 7, where the identifying parameters in the parameter set ~~that do not contribute to content~~ includes:

retrieving content corresponding to a sampled URL including a combination of parameters in the parameter set; and

identifying the combination of parameters as corresponding to retrieved content, where the retrieved content is approximately the same as another retrieved content corresponding to another combination of parameters that includes a reduced number of parameters.

10. (previously presented) The method of claim 9, where the combination of parameters includes at least one of the sampled URL with no parameters, the sampled URL with individual parameters, or the sampled URL with combinations of the at least one parameter.

11. (previously presented) The method of claim 7, where the rewrite rule applies to a particular web site or web host.

12. (Currently amended) One or more devices comprising:
at least one fetch bot ~~configured~~ to download content on a network from locations specified by uniform resource locators (URLs);
a content manager configured to extract URLs from the downloaded content;
a rewrite component ~~configured to~~

receive a URL that refers to content and that includes a parameter set including at least one parameter,

apply a predetermined rewrite rule to the URL that removes the at least one parameter from the URL when the at least one parameter does not affect the content referred to by the URL, where the predetermined rewrite rule is determined by receiving a plurality of URLs that include the parameter set and identifying parameters in the parameter set that do not ~~contribute to~~ result in retrieving substantially different content, when present in a URL; and

output the rewritten URL as the canonical form of the URL; and
a URL manager configured to store the canonical form of the URL.

13. (Cancelled)

14. (Currently amended) The one or more devices of claim 12, where the identifying parameters in the parameter set ~~that do not contribute to content~~ includes:

retrieving content corresponding to a sampled URL including a combination of parameters in the parameter set; and

identifying the combination of parameters as corresponding to retrieved content, where the retrieved content is approximately the same as another retrieved content corresponding to another combination of parameters that includes a reduced number of parameters.

15. (previously presented) The one or more devices of claim 14, where the combination of parameters includes at least one of the sampled URL with no parameters, the sampled URL with individual parameters, or the sampled URL with combinations of the at least one parameter.

16. (previously presented) The one or more devices of claim 12, where each rewrite rule applies to a particular web site or web host.

17. (Currently amended) A system comprising:

- means for receiving a first uniform resource locator (URL) including one or more parameters;
- means for retrieving content corresponding to the first URL;
- means for retrieving content corresponding to a plurality of URLs having different parameter combinations of the one or more parameters;
- means for identifying the parameter combination from the plurality of URLs that corresponds to content that is approximately the same as the content corresponding to the first URL and that contains a minimum number of parameters compared to other parameter combinations; and
- means for generating one or more URL rewrite rules based on the identified parameter combination.

18. (previously presented) A computer-readable memory device including programming instructions executed by a processor, the programming instructions comprising:

instructions for receiving a first uniform resource locator (URL) including one or more parameters;

instructions for retrieving content corresponding to the first URL;

instructions for retrieving content corresponding to a plurality of URLs having different parameter combinations of the one or more parameters;

instructions for identifying the parameter combination from the plurality of URLs that corresponds to content that is approximately the same as the content corresponding to the first URL and that includes a minimum number of parameters; and

instructions for generating one or more URL rewrite rules based on the identified parameter combination.

19. (previously presented) The system of claim 17, where the parameter combination comprises one of the first URL with no parameters, the first URL with each of the one or more parameters individually, or the first URL with combinations of the one or more parameters.

20. (previously presented) The computer-readable memory device of claim 18, where the instructions for receiving a first URL, the instructions for retrieving content corresponding to the first URL, the instructions for retrieving content corresponding to a plurality of URLs, and the instructions for identifying the parameter combination are

performed for multiple first URLs, each first URL including the one or more parameters, and where the one or more URL rewrite rules specify that parameters that do not occur in a threshold number of the identified parameter combinations are to be removed.

21. (New) The system of claim 17, further comprising:

means for determining whether the content that corresponds to the plurality of URLs is approximately the same as the content that corresponds to the first URL using a similarity hash function.

22. (New) The computer-readable memory device of claim 18, where the rewrite rules specify that parameters that do not occur in a threshold number of the identified parameter combinations are to be removed.